Lareena Llamado

**Kickstarter Data Report**

1. **Given the provided data, what are three conclusions we can draw about Kickstarter campaigns?**

Three conclusions from provided data is as follows.

Top 3 categories that were successful are theater, music, and film and video. Top 3 sub-categories are plays, rock, and wearables.

Top 3 successful months were May, June, and February.

Most successful country was the United States.

1. **What are some limitations of this dataset?**

A limitation of this dataset is that it only classified 21 countries. To get a full picture of Kickstarter projects throughout the world you could argue that all 270 countries should be included in the dataset.

Another limitation is the number of categories. It seems small to establish what is and is not successful with Kickstarter projects. As well as limited number of sub-categories.

1. **What are some other possible tables and/or graphs that we could create?**

Possible tables that could be created could be the percentage of countries within each category. This could possibly show what categories are succeeding in which countries. Another graph that could be created is a pivot table of time of conversion versus successful and failed categories and another pivot table of time of conversion versus successful and failed categories. This would show further what time frame each endeavor took to be successful within its category and sub-category. It would also show if time is a factor when creating a Kickstarter project.

**Bonus Statistical Analysis**

1. **Use your data to deter whether the mean or the median summaries the data more meaningfully.**

The mean of the backers summaries the data more meaningfully. The mean of successful versus failed projects, 194 to 18, respectively. It shows that the greater the mean the probability of the project succeeding. More backers mean more money which means the likelihood of the project succeeding as opposed to failing.

1. **Use your data to determine if there is more variability with successful or unsuccessful campaigns. Does this make sense? Why or why not?**

There is more variability with successful campaigns as opposed to failed campaigns. The variance of the data shows how far the values from the data set are from the means. The variance for successful and failed are 713,167 and 5,331, respectively. This shows that the values for successful campaigns vastly differ in how many backers there were for different campaigns. The more backers a campaign has the more likely it will be successful.

The variability of the data makes sense because it shows that overall the successful campaigns did have a substantial amount of backers compared to the failed campaigns. The disparity of the between backers on successful campaigns as opposed to backers on failed campaigns shows that more backers may provide more monetary support and thus helping the campaign succeed.